상대적 소속 함수에 기반을 둔 새로운 유사성 척도와 언어 근사에의 응용

최 대 영[†]

요 약

상대적 소속 함수(RMF)에 기반을 둔 새로운 유사성 척도를 제안한다. 본 논문에서 RMF는 퍼지 부분 집합간의 상대성을 쉽게 나타내기 위해 제시되었다. 이러한 RMF의 형태는 매개변수값들에 따라 결정되기 때문에 매개변수 값들만을 조정해 줌으로써 퍼지 부분 집합간의 상대성을 쉽게 나타낼 수 있다. 그러므로 퍼지 부분 집합을 이용해 주관성을 표현할 때 개인이나 문화차이간의 상대성을 쉽게 반영해 줄 수 있다. 이 경우 이들 매개변수들은 퍼지 부분 집합의 구조를 결정해 주는 특징점들이라고 할 수 있다. 결과적으로 퍼지 부분 집합간의 유사성 정도가 RMF의 매개변수들을 이용해서 빠르게 계산될 수 있다. RMF에 의해 표현된 퍼지 부분 집합간의 유사성 정도를 계산하기 위해 유클리디안 거리를 사용한다. 한편, 제안된 유사성 척도의 응용 분야로 새로운 언어 근사 방법을 제시하고 수치적인 예를 보여준다.

A New Similarity Measure based on RMF and Its Application to Linguistic Approximation

Dae-Young Choi[†]

ABSTRACT

We propose a new similarity measure based on relative membership function (RMF). In this paper, the RMF is suggested to represent the relativity between fuzzy subsets easily. Since the shape of the RMF is determined according to the values of its parameters, we can easily represent the relativity between fuzzy subsets by adjusting only the values of its parameters. Hence, we can easily reflect the relativity among individuals or cultural differences when we represent the subjectivity by using the fuzzy subsets. In this case, these parameters may be regarded as feature points for determining the structure of a fuzzy subset. In the sequel, the degree of similarity between fuzzy subsets can be quickly computed by using the parameters of the RMF. We use Euclidean distance to compute the degree of similarity between fuzzy subsets represented by the RMF. In the meantime, we present a new linguistic approximation method as an application area of the proposed similarity measure and show its numerical example.

키워드 : 유사성 척도, 상대적 소속 함수, 언어 근사

1. 서 론

프로그래밍 언어에서 변수가 같은 자료형의 다양한 값들을 가질 수 있는 것처럼 인간이 사용하는 언어에서도 언어 변수(Linguistic Variable)[10]가 그에 관련된 다양한 언어 값(Linguistic Value)들을 가질 수 있다. 예를 들어, "나이가 젊다" 에서 '나이'가 언어 변수가 되고 '젊다'가 언어 값이 된다. 일반적으로 이러한 언어 값들은 퍼지 부분 집합의 개념에 의해 표현된다[10]. 이와 같은 언어 값을 사용하는 개인간에 그 언어 값에 대한 개념적인 차이가 명확히 존재하지만 개인간의 비교를 위한 표준을 만들기 위해서는 표준화된 소속 함수(Standardized Membership Function)의 사용이 요구된다[13]. 이러한 관점에서 본 논문은 상대적 소속 함수(Relative Membership Function ; RMF)를 제안한다. 이러한 RMF는 언어 값을 사용하는 개인간에 그 언어 값에 대한 개념적인 차이인 상대성을 쉽게 나타내기 위해 제시되었다. RMF의 형태는 매개변수 값들에 따라 결정되기 때문에 매개변수 값들만을 조정해 줌으로써 퍼지 부분 집합간의 상대성을 쉽게 나타낼 수 있다. 이 경우 이들 매개변수들은 퍼지 부분 집합의 구조를 결정 해주는 특징점들 이라고 할 수 있다. 결과적으로 퍼지 부분 집합간의 유사성 정도가 RMF의 매개변수들을 이용해서 빠르게 계산될 수 있는 장점을 제공한다. 본 논문에서는 RMF에 의해 표현된 퍼지 부분 집합간의 유사성 정도를 계산하기 위해 유클리디안 거리를 사용한다. 본 논문은 2장에서 RMF를 제안하고 그에 관련된 특성을 분석한다.

※ 이 논문은 한국과학재단의 해외 post-doc. 연수지원비 및 2000년도 유한대학 학술연구비 지원에 의하여 연구되었음.
† 정 회 원 : 유한대학 경영정보과 교수
논문접수 : 2001년 7월 23일, 심사완료 : 2001년 9월 21일

3장에서 RMF에 기반을 둔 유사성 측도를 제안하고, 4장에서는 제안된 유사성 측도를 이용한 새로운 형태의 언어 근사 방법을 제시하고 수치적인 예를 보여준다. 끝으로 5장에서 결론 및 추후 연구에 대해 논한다.

2. RMF(Relative Membership Function)

소속 함수는 퍼지 집합의 중요한 구성 요소이다. 이러한 결과로 다양한 형태의 소속 함수들이 제시되었다[2-7, 12, 13]. 그러나 기존의 소속 함수는 개인, 국가, 문화간의 차이로 인해 발생하는 소속 함수들간의 상대성을 반영해 주려는 연구가 미약하였다. 그러한 결과로 기존의 소속 함수는 일반적으로 임의적인(Ad-hoc) 방법을 사용해서 개별적으로 개발되었다. 이 경우 만약 많은 사람들의 주관성을 표현해 주기 위해 여러 개의 소속 함수가 필요하다면 얼마나 많은 소속 함수들이 필요한가에 대한 질문이 제기된다. 만약 퍼지 시스템을 구현할 때 적은 개수의 소속 함수 사용을 원한다면 소속 함수는 특정 상황에 적응가능해야 한다. 이는 조정 가능한 매개변수들을 사용해서 해결할 수 있다. 이러한 문제점을 다루기 위해 매개변수들을 사용하는 RMF를 제안한다. RMF의 매개변수 값들은 개인, 국가, 문화등의 차이에 따라 다르게 얻어진다. RMF의 형태는 매개변수 값들에 따라 결정되기 때문에 매개변수 값들만을 조정해 줌으로써 퍼지 부분 집합간의 상대성을 쉽게 나타낼 수 있다. 이 경우 이들 매개변수들은 퍼지 부분 집합의 구조를 결정해주는 특징점들이라고 할 수 있다. 결과적으로 퍼지 부분 집합간의 유사성 정도가 RMF의 매개변수들을 이용해서 빠르게 계산될 수 있는 장점을 제공한다. 비록 RMF가 형태에 있어 삼각형 타입, 사다리꼴 타입등과 유사하지만, 본 논문에서는 상대성을 쉽게 반영해 주는 새로운 소속 함수를 제시하였고, 매개변수 값들을 어떻게 결정하는지를 명확히 정의하였다. 이러한 RMF는 구조가 단순하여 사용자가 다루기 쉬운 장점을 가지고 있다.

2.1 RMF의 형태

RMF는 포인트 타입 RMF(Point-type RMF)와 구간 타입 RMF(Interval-type RMF) 가 있다.

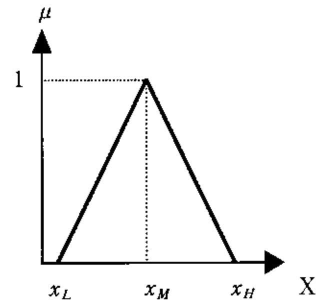
2.1.1 포인트 타입 RMF

A가 어떤 언어 값을 위한 퍼지 부분 집합이라 하고 전체 집합 X의 부분집합이라 하자. 그러면, $x \in X$에 대해 포인트 타입 RMF를 다음과 같이 정의할 수 있다.

$$\mu_A(x, x_M, x_L, x_H) = 1 - [|(x - x_M)/(x_H - x_L)| \times 2], \quad (1)$$

단, $x_L \le x \le x_H$ 이고 $|x_M - x_L| = |x_H - x_M|$이다. 만약 소속 함수 (1)의 결과가 [0, 1]로 정규화 되어 있다면, $x \in (-\infty, x_L] \cup [x_H, \infty)$]에서 $\mu_A(x, x_M, x_L, x_H) = 0$이다.

식 (1)에서 x_M는 x_L과 x_H의 중앙점이고 퍼지 부분 집합 A에 절대적으로 포함되는 퍼지 부분 집합 A의 Kernel이라고 가정한다. 결과적으로 포인트 타입 RMF는 확실성이 한 점일 때 사용된다. 또한, $[x_L, x_H]$이 퍼지 부분 집합 A의 가능구간이라고 가정한다. 이때 x_M과 $[x_L, x_H]$는 개인이나 문화 차이 등에 따라 다르게 결정된다. 식 (1)은 x_M에 대해 대칭성을 갖으며 만약 소속 함수 (1)의 결과가 [0, 1]로 정규화 되어있다면 다음과 같은 형태가 된다.



(그림 1) 포인트 타입 RMF의 형태

[예 1] 퍼지 부분 집합 A를 '*10에 가까운 실수*'라고 하면 포인트 타입 RMF를 사용한 그에 대한 소속 함수는 다음과 같이 정의될 수도 있다.

$$\mu_A(x, x_M, x_L, x_H) = \mu_A(x, 10, 5, 15)$$
$$= 1 - [|(x-10)/(15-5)| \times 2]$$

쉬운 예를 들기 위해 x가 정수이고 소속 함수의 계산 결과가 소숫점 아래 한자리까지만 표현될 때의 실험 결과를 보여준다.

〈표 1〉 포인트 타입 RMF의 예

x	소 속 함 수 값
5	0.0
6	0.2
7	0.4
8	0.6
9	0.8
10	**1.0**
11	0.8
12	0.6
13	0.4
14	0.2
15	0.0

만약 x를 위해 실수가 사용되었거나 소속 함수의 계산 결과가 더 많은 소숫점 아래 자리수까지 표현되었다면 <표 1>에 있는 계산 결과의 중간값들이 얻어진다. 또한, 이 경우 중앙점인 10에대해 소속 함수 값들이 대칭성을 갖는 것을 알 수 있다.

2.1.2 구간 타입 RMF

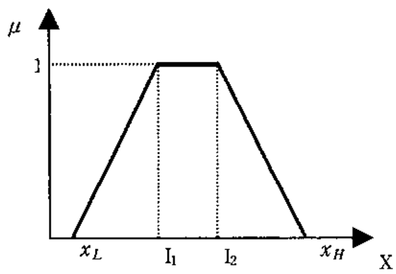
A가 어떤 언어 값을 위한 퍼지 부분 집합이라 하고 전체

집합 X의 부분 집합이라 하자. 그러면, $x \in X$에 대해 구간 타입 RMF를 다음과 같이 정의할 수 있다.

$$\mu_A(x, I_1, I_2, x_L, x_H) = \begin{cases} 1-[|(x-I_1)/(x_H-x_L)|\times 2] & \text{for } x_L \le x < I_1 \\ 1 & \text{for } I_1 \le x \le I_2 \\ 1-[|(x-I_2)/(x_H-x_L)|\times 2] & \text{for } I_2 < x \le x_H \end{cases} \quad (2)$$

단, $x_L \le x \le x_H$이고 $|I_1 - x_L| = |x_H - I_2|$이다. 만약 소속 함수 (2)의 결과가 [0, 1]로 정규화되어 있다면, $x \in (-\infty, x_L] \cup [x_H, \infty)$에서 $\mu_A(x, I_1, I_2, x_L, x_H) = 0$이다.

식 (2)에서 확실성 구간인 [I_1, I_2]는 퍼지 부분 집합 A에 절대적으로 포함되는 퍼지 부분 집합 A의 Kernel이라고 가정한다. 결과적으로 구간 타입 RMF는 확실성이 구간으로 주어질 때 사용된다. 또한, $[x_L, x_H]$이 퍼지 부분집합 A의 가능 구간이라고 가정한다. 이 때 [I_1, I_2]과 $[x_L, x_H]$는 개인이나 문화 차이등에 따라 다르게 결정된다. 식 (2)는 구간 [I_1, I_2]에 대해 대칭성을 갖으며 만약 소속 함수 (2)의 결과가 [0, 1]로 정규화되어 있다면 다음과 같은 형태가 된다.



(그림 2) 구간 타입 RMF의 형태

[예 2] 퍼지 부분집합 A를 '*젊은 남자*'라고 하면 구간 타입 RMF를 사용한 그에 대한 소속 함수는 다음과 같이 정의될 수도 있다.

$$\mu_A(x, I_1, I_2, x_L, x_H) = \mu_A(x, 20, 30, 15, 35)$$

$$= \begin{cases} 1-[|(x-20)/(35-15)|\times 2] & \text{for } 15 \le x < 20 \\ 1 & \text{for } 20 \le x \le 30 \\ 1-[|(x-30)/(35-15)|\times 2] & \text{for } 30 < x \le 35 \end{cases}$$

쉬운 예를 들기 위해 x가 정수이고 소속 함수의 계산 결과가 소숫점 아래 한자리까지만 표현될 때의 실험결과를 보여준다.

만약 x를 위해 실수가 사용되었거나 소속 함수의 계산결과가 더 많은 소숫점 아래 자리수까지 표현되었다면 <표 2>에 있는 계산결과의 중간값들이 얻어진다. 이 경우 구간 [20, 30]에 대해 소속 함수 값들이 대칭성을 갖는 것을 알 수 있고, 또한 <표 2>는 좌측 끝점(x_L)과 우측 끝점(x_H)에서 소속 함수의 계산 결과가 0이 아닌 것을 알 수 있다. 이는 소속 함수가 모호성을 모델링하기 위해서 제안되었기 때문에 소속 함수의 지나친 정확성은 모호성 표현에 대해 모순이 된다는 관점에서 자연스럽다[4]. 이러한 관점에서 두 가지 형태의 RMF에 대해 좌측 끝점(x_L)과 우측 끝점(x_H)에서 소속 함수의 계산 결과가 0이 아닌 것이 허용된다. 이 경우 RMF의 결과는 α-레벨 집합과 유사하다(단, $\alpha \in [0, 1]$). 만약 이러한 소속 함수의 계산 결과를 구간 [0, 1]로 정규화해 줄 필요가 있다면 다음의 [예 3]과같이 약간의 조정을 해주면 된다.

〈표 2〉 구간 타입 RMF의 예

x	소 속 함 수 값
15	0.5
16	0.6
17	0.7
18	0.8
19	0.9
[20, 30]	**1.0**
31	0.9
32	0.8
33	0.7
34	0.6
35	0.5

[예 3] <표 2>에서 소속 함수 값들을 구간 [0, 1]로 정규화해 줄 필요가 있을 경우, 즉 [0.5, 1]에서 [0, 1]으로의 정규화를 하기 위해 첫째, [0.5, 1]에서 [0, 0.5]으로의 매핑 함수가 수행된다. 즉, {0.5, 0.6, ⋯, 0.9, 1.0}에 있는 각 원소에서 0.5를 뺀다. 둘째, 그 결과인 {0.0, 0.1, ⋯, 0.4, 0.5}의 각 원소에 대해 2를 곱해줌으로써 구간 [0, 1]으로의 정규화가 얻어진다.

2.2 RMF의 특성

RMF는 다음과 같은 특성들을 갖는다. 첫째, RMF는 포인트 타입은 x_M에 대해 그리고 구간 타입은 [I_1, I_2]에 대해 대칭성을 갖는다. 둘째, 어떤 언어 값의 가능 구간이 $[x_L, x_H]$일 때, RMF의 계산은 포인트 타입은 식 (1)에 따라 x_M으로부터의 거리에 대한 함수이고, 구간 타입은 식 (2)에 따라 [I_1, I_2]으로부터의 거리에 대한 함수로 구해진다. 셋째, RMF는 구조가 단순하여 사용자가 다루기 쉬운 장점을 가지고 있다. 넷째, 매개변수값들을 조정해 줌으로써 퍼지 부분 집합간의 상대성을 적응적으로 표현할 수 있다. 이는 RMF의 형태가 매개변수들의 값에 따라 적응적으로 변화되기 때문이다. 결과적으로 RMF는 매개변수 값들을 조정해 줌으로써 상황에 적응할 수 있다. 이러한 관점에서 RMF를 사용하는 퍼지 시스템은 개인, 국가, 문화 등의 차이로 인한 소속 함수간의 상대성을 쉽게 반영해 줄 수 있다.

2.3 RMF와 기존 소속 함수의 비교

기존의 소속 함수들에서는 개인, 국가, 문화 등의 차이로

인해 발생하는 소속 함수간의 상대성을 반영해주려는 연구가 미약하였다. 그러한 결과로 기존의 소속 함수는 일반적으로 임의적인 방법을 사용해서 개별적으로 개발되었다. 결과적으로 퍼지 부분 집합간의 상대성을 계산하기 어려운 문제점을 가지고 있다. 이러한 문제를 다루기 위해 매개변수들을 사용하는 RMF를 제안하였다. RMF에서는 개인, 국가, 문화 등의 차이에 따라 매개변수값들이 다르게 구해지고 이러한 특성을 이용해 그러한 차이로 인해 발생하는 퍼지 부분 집합간의 상대성을 쉽게 나타내 줄 수 있다. 이 경우 이들 매개변수들은 퍼지 부분 집합의 구조를 결정해 주는 특징점들 이라고 할 수 있다. 결과적으로 퍼지 부분 집합간의 유사성 정도가 RMF의 매개변수들을 이용해서 빠르게 계산될 수 있는 장점을 제공한다. 기존의 소속 함수들은 일반적으로 복잡한 수식을 사용해서 표현되기 때문에 실제 응용분야에 적용될 때 사용하기 어려운 문제점이 있다. 이에 반해, RMF는 구조가 단순하여 사용자가 다루기 쉬운 장점을 가지고 있다. 한편 퍼지 집합 이론을 전문가 시스템, 퍼지 제어기 등에 응용할 때 중요한 고려사항은 계산효율성이다. RMF는 식 (1)과 (2)에서와 같은 선형식을 사용하기 때문에 상대적으로 계산 복잡도가 줄어들게 된다. 예를 들어 Π타입 소속 함수는 이차식을 사용하고, 다항식[5, 7]이나 지수 함수[2, 4, 12]를 이용하는 경우도 있다.

〈표 3〉 RMF와 기존 소속 함수의 비교

비교항목	기존 소속 함수	RMF
상 대 성	반영하기 어려움	반영하기 쉬움
적 응 성	제 한 적	좋 음
형 태	임 의 적	포인트 타입 또는 구간 타입
계산복잡도	상대적으로 복잡	선형(Linear)

3. RMF에 기반을 둔 유사성 측도(Similarity Measure)

유사성 측도(SM)는 퍼지 부분 집합간의 유사성 정도를 표현하기 위한 방법의 하나로 퍼지 시스템에서 퍼지 부분 집합을 다룰 때 자주 사용되는 개념이다. Turksen과 Zhong[8]은 두 퍼지 집합간의 유사성 측도를 구할 때 다음과 같이 상이성 측도(Distance Measure ; DM)를 이용하여 구했다. $SM=(1+DM)^{-1}$, (단, $SM \in [0, 1]$). 그러나 그들은 DM를 어떻게 구하는지 명확히 제시하고 있지 않다. Zwick[15]등은 인간의 행동적인 실험에 기반을 두고 5가지 좋은 DM를 제시했는데, 그 실험 결과는 인간이 퍼지한 개념을 처리하거나 결합할 때 복잡한 소속함수들을 특징적인 측면 위주로 단순화하여 해결한다는 결론을 얻었다. 이들은 단순화의 수단으로 최소값, 최대값, α-레벨 집합 등을 사용하였다. 이러한 실험 결과에 기반을 두고 본 논문에서는 RMF의 매개변수들을 이용하여 소속 함수를 단순화시켰고, 본 논문에서 제시한 유사성 측도는 이러한 RMF에 기반을 두고 설계되었다. 3장에서는 퍼지 부분 집합들이 식 (2)에 있는 구간 타입 RMF를 가지고 표현되어 있을 경우에 대한 퍼지 부분 집합간의 유사성 측도를 제시였는데 포인트 타입 RMF인 경우에도 비슷한 방법으로 계산될 수 있다.

두 개의 퍼지 부분 집합이 구간 타입 RMF를 사용해서 각각 $(I_1^1, I_2^1, x_L^1, x_H^1)_1 = T_1$,와 $(I_1^2, I_2^2, x_L^2, x_H^2)_2 = T_2$로 정의되었다고 하자. 이 경우 윗첨자는 T_1과 T_2를 구별하기 위한 것이다. 이렇게 표현된 두 개의 퍼지 부분 집합과 $(I_1^{'}, I_2^{'}, x_L^{'}, x_H^{'}) = T'$와의 유사성 정도를 나타내는 유사성 측도를 다음과 같이 유클리디안 거리를 사용해 구할 수 있다.

$$d_1(T_1, T') = \sqrt{(I_1^1 - I_1^{'})^2 + (I_2^1 - I_2^{'})^2 + (X_L^1 - X_L^{'})^2 + (X_H^1 - X_H^{'})^2}$$

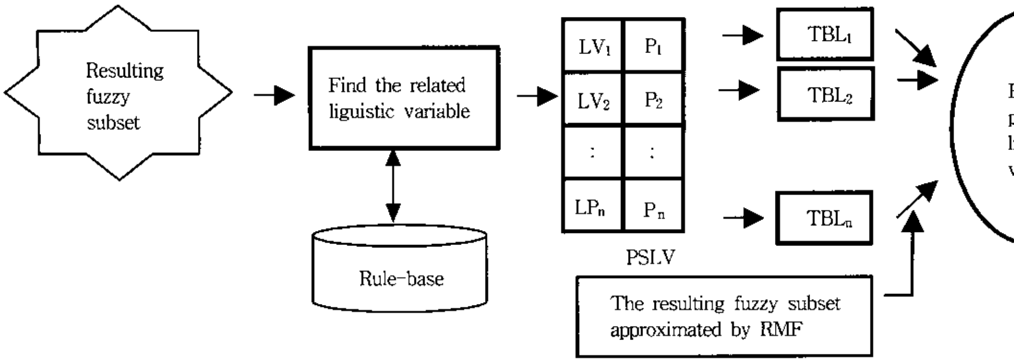
$$d_2(T_2, T') = \sqrt{(I_1^2 - I_1^{'})^2 + (I_2^2 - I_2^{'})^2 + (X_L^2 - X_L^{'})^2 + (X_H^2 - X_H^{'})^2}$$

이 경우 만약 $d_1 > d_2$가 된다면 T_2가 T_1보다 T'에 상대적으로 유사하다고 할 수 있다. RMF는 소속 함수들을 단순화 시키는 방법으로 매개변수들을 이용하였고, 본 논문에서 제안된 유사성 측도는 이러한 매개변수를 이용해 구할 수 있다. R. Degani와 G. Bortolan[16]에 의하면 언어 근사를 구하는데 있어 유클리디안거리가 최적의 언어 값을 구하는 방법으로 제시되었다. 다만 이 경우 적절한 처리 시간을 얻기 위해서는 처리 대상 언어 값들의 갯수가 적어야 한다[16]. 본 논문의 4장에서 제안된 언어 근사 방법은 퍼지 시스템에서 사용되는 언어 변수들을 분리하는 분할 개념을 적용했기 때문에 각각의 언어 변수는 적은 갯수의 언어 값들로 구성된다((그림 3) 참조). 이러한 관점에서 유사성 측도를 구하는 방법으로 유클리디안거리를 사용하였다.

4. 언어 근사(Linguistic Approximation)에의 응용

언어 근사는 인간과 기계간에 상호작용이 있는 응용 분야에서 매우 중요하다. 그러나, 이는 쉬운 문제가 아니기 때문에 아직 연구가 미약한 실정이다[1, 9]. 본 논문에서 제안된 유사성 측도의 전형적인 응용 분야로 RMF에 기반을 둔 새로운 형태의 효율적인 언어 근사 방법을 제시한다.

4장에서는 퍼지 시스템의 계산 결과로 만들어진 퍼지 부분 집합과 가장 유사한 언어 값의 퍼지 부분 집합을 발견하려 한다. 그러한 언어 값을 언어 근사되었다고 한다. 어떤 퍼지 시스템에 대해 가나다 순서로 정렬된 미리 정의된 언어 변수들의 집합이 있다고 하고 이를 PSLV라 하자. 이러한 PSLV에 있는 각 언어 변수는 그에 관련된 언어 값들로 구성된 테이블을 가리키는 포인터가 있다. 즉, PSLV는 각 언어 변수에 관련된 언어 값들을 빠르게 발견하기 위한 인덱스 테이블이라고 할 수 있다. 이와같이 언어 근사 과정에서 어떤 언어 변수에 관련 없는 언어 값들에 대한 탐색을 방지하기 위해 퍼지 시스템에서 사용되는 언어 변수들을 분리하



(그림 3) 언어 근사의 과정

는 분할 개념을 사용하였다. 결과적으로 각 테이블은 해당 언어 변수에 관련된 언어 값들로 구성된다. 한편, 언어 변수에 관련된 각 테이블에 있는 언어 값들은 2장에서 제시한 RMF에 의해 정의되어 있다고 가정한다.

퍼지 시스템의 계산 결과로부터 만들어진 퍼지 부분 집합은 그와 관련된 언어 변수로 쉽게 연관시킬 수 있다. 예를 들어 주어진 규칙 '*If* X is a *then* Y is b'에서 언어 값 'a'는 언어 변수 'X'와 연관되는 반면 언어 값 'b'는 언어 변수 'Y'와 연관된다. 그러므로 규칙 베이스에서 선택된 규칙을 참조하면 쉽게 연관된 언어 변수를 발견할 수 있다. PSLV에서 연관된 언어 변수(LV)를 결정한 후에, 그 언어 변수의 포인터 (P_i, $i=1, 2, \cdots, n$)를 사용해서 그에 대응하는 테이블(TBL)이 결정된다. 이렇게 결정된 테이블에 있는 RMF에 의해 표현된 관련된 퍼지 부분 집합들에 대해 유클리디안 거리를 사용해서 테이블에 있는 관련된 퍼지 부분 집합들과 RMF에 의해 근사된 퍼지 시스템의 계산 결과간의 거리들이 각각 계산된다. 결과적으로 가장 적은 거리를 갖는, 즉 퍼지시스템의 계산 결과와 가장 유사한 언어 값이 언어 근사로 선택된다. 결과적으로 본 논문에서 제안된 언어 근사 방법은 2장에서 제시된 RMF의 매개변수들을 이용하여 계산된다. 이 경우 매개변수들은 언어 값들을 나타내기 위한 특징점들 이라고 할 수 있다. 결과적으로 언어 근사를 위한 계산 복잡도가[1, 9] 등과 비교할 때 상대적으로 줄어든다. 왜냐하면 퍼지 부분 집합에 있는 모든 원소에 대해 비교를 해야 하는 NPMF(Non-Parametric Membership Function)에 기반을 둔 언어 근사 방법과 비교할 때 RMF에 기반을 둔 언어 근사는 매개변수만을 비교해서 계산이 효율적으로 되기 때문이다.

[예 4] 언어 변수 '나이'에 관한 테이블이 '젊다'와 '매우 젊다'라는 두개의 언어 값으로 구성되어 있다고 가정하자. 이때 이들 언어 값들이 식 (2)에 있는 구간 타입 RMF를 사용해서 각각 (I_1, I_2, x_L, x_H, 언어값$)_1$ =(20, 30, 15, 35, '젊다'$)_1$ = T_1와 (I_1, I_2, x_L, x_H, 언어값$)_2$ =(20, 27, 17, 30, '매우 젊다'$)_2$ = T_2로 정의되어 있다고 하자. 또한, 퍼지 시스템의 계산 결과인 퍼지 부분 집합이 제안된 RMF 형태로 다음과 같이 (I_1', I_2', x_L', x_H') = (20, 25, 16, 29) = T′로 매개변수화되어 있다고 하자. 제안된 유사성 척도를 사용해서 언어 근사를 구할 수 있다.

$$d_1(T_1, T') = \sqrt{(I_1-I_1')^2+(I_2-I_2')^2+(X_L-X_L')^2+(X_H-X_H')^2}$$
$$= \sqrt{(20-20)^2+(30-25)^2+(15-16)^2+(35-29)^2} = \sqrt{62},$$
$$d_2(T_2, T') = \sqrt{(I_1-I_1')^2+(I_2-I_2')^2+(X_L-X_L')^2+(X_H-X_H')^2}$$
$$= \sqrt{(20-20)^2+(27-25)^2+(17-16)^2+(30-29)^2} = \sqrt{6}.$$

이 경우 퍼지 시스템의 계산 결과인 T′에 대해 $d_1 > d_2$이기 때문에 언어 근사로 언어 값 '매우 젊다'가 선택된다. 비슷한 방법으로 m개의 언어 값들로 구성된 테이블인 경우 제안된 유사성 척도를 반복 적용함에 의해 (d_1, d_2, $\cdots$, d_m)를 계산하고 min(d_1, d_2, $\cdots$, d_m)인 언어 값이 언어 근사로 선택된다. 어떤 언어 변수에 대해 해당 테이블이 더 많은 언어 값들을 가지고 있으면 보다 정확한 언어 근사를 구할 수 있다. 한편, 퍼지 집합 이론을 전문가 시스템, 자연어 처리, 퍼지 제어기 등에 적용할 때 중요한 고려사항은 상대적으로 빠른 방법을 고안해 내는 것이다[10, 11, 14]. 이러한 공학적인 관점에서 본 논문에서 제안된 언어 근사 방법은 이전의 연구[1, 9]들과 비교할 때 처리 속도 측면에서 장점을 갖는다. 이는 RMF를 갖는 퍼지 시스템이 NPMF를 갖는 퍼지 시스템보다 매개변수를 이용하여 상대적으로 빠르게 처리될 수 있고, 또한 언어 근사 과정에서 관련 없는 언어 값들에 대한 탐색을 방지하기 위해 퍼지 시스템에서 사용되는 언어 변수들을 분리하는 분할 개념을 적용했기 때문이다.

5. 결론 및 추후 연구

언어 값을 사용하는 개인간에 그 언어 값에 대한 개념적인 차이가 명확히 존재하지만 개인간의 비교를 위한 표준을 만들기 위해서는 표준화된 소속 함수의 사용이 요구된다[13]. 이러한 관점에서 본 논문은 RMF를 제안하였다. 이러한 RMF는 언어 값을 사용하는 개인간의 그 언어 값에 대한 개념적인 차이인 상대성을 쉽게 나타내기 위해 제시되었다. RMF의

형태는 매개변수 값들에 따라 결정되기 때문에 매개변수 값들만을 조정해 줌으로써 퍼지 부분 집합간의 상대성을 쉽게 나타낼 수 있다. 이 경우 이들 매개변수들은 퍼지 부분 집합의 구조를 결정 해주는 특징점들 이라고 할 수 있다. 결과적으로 퍼지 부분 집합간의 유사성 정도가 RMF의 매개변수들을 이용해서 빠르게 계산될 수 있는 장점을 제공한다. 또한, 본 논문에서는 RMF에 의해 나타내진 퍼지 부분 집합간의 유사성 정도를 계산하기 위해 유사성 측도를 제안하였다. 이러한 유사성 측도를 이용한 새로운 형태의 언어 근사 방법을 제안하였다. 본 논문에서 제안된 언어 근사 방법은 RMF의 매개변수들을 사용해서 구해지고, 또한 언어 근사 과정에서 관련 없는 언어 값들에 대한 탐색을 방지하기 위해 퍼지 시스템에서 사용되는 언어 변수들을 분리하는 분할 개념을 적용했기 때문에 이전의 연구[1, 9]들과 비교할 때 처리 속도 측면에서 장점이 있다.

본 논문과 관련된 추후 연구는 첫째, 본 논문에서 제안된 RMF는 선형적으로 표현되기 때문에 정교한 소속 함수를 표현하기 어렵다. 이러한 관점에서 비선형적으로 표현된 정교한 소속 함수를 매개변수를 갖는 구간별 선형 소속 함수로 근사 변환해 주는 방법이 필요하다. 둘째, (그림 3)과 [예 4]에서 퍼지 시스템의 계산 결과인 퍼지 부분 집합을 제안된 RMF 형태로 근사 변환해 주는 연구가 필요하다. 셋째, 제안된 RMF의 일반성을 강화하기 위해 비대칭성을 갖는 RMF에 대한 연구와 보다 더 정교한 소속 함수를 표현할 수 있도록 보다 더 많은 매개 변수를 갖는 확장된 RMF에 대한 연구가 필요하다.

참 고 문 헌

[1] F. Eshragh and E. H. Mamdani, "A General Approach to Linguistic Approximation," *International Journal of Man-Machine Studies*, Vol.11, pp.501-519, 1979.

[2] A. Grauel and L. A. Ludwig, "Construction of differential membership functions," *Fuzzy Sets and Systems*, Vol.101 pp.219-225, 1999.

[3] G. J. Klir and T. A. Folger, *Fuzzy sets, uncertainty and information*, Prentice Hall, 1988.

[4] V. Novak, *Fuzzy sets and their applications*, Adam Hilger, 1989.

[5] W. Pedrycz, *Fuzzy control and fuzzy systems*, Research Studies Press, 1989.

[6] W. Pedrycz, "Fuzzy set technology in knowledge discovery," *Fuzzy sets and systems*, Vol.98, pp.279-290, 1998.

[7] T. A. Runkler and J. C. Bezdek, "Function Approximation with Polynomial Membership Functions and Alternation Cluster Estimation," *Fuzzy Sets and Systems*, Vol.101, pp.207-218, 1999.

[8] I. B. Turksen and Zhao Zhong, "An Approximate analogical Reasoning Approach based on Similarity Measures," IEEE Trans. on SMC, Vol.18, No.6, pp.1049-1056, 1988.

[9] F. Wenstop, "Deductive Verbal Models of Organization," *International Journal of Man-Machine Studies*, Vol.8, pp. 293-311, 1976.

[10] L.A. Zadeh, "The Concept of A Linguistic Variable and Its Application to Approximate Reasoning - 1," in : R. R. Yager, S. Ovchinnikov, R. M. Tong and H. T. Nguyen (Eds.), Fuzzy sets and Applications : *Selected papers by L. A. Zadeh*, John Wiley & Sons, pp.219-269, 1987.

[11] ________, "A Theory of Approximate Reasoning," in : R. R. Yager, S. Ovchinnikov, R. M. Tong and H. T. Nguyen (Eds.), Fuzzy sets and Applications : *Selected papers by L. A. Zadeh*, John Wiley & Sons, pp.367-412, 1987.

[12] ________, "A Fuzzy-set-theoretic Interpretation of Linguistic Hedges," in : R. R. Yager, S. Ovchinnikov, R. M. Tong and H. T. Nguyen (Eds.), Fuzzy sets and Applications : *Selected pa- pers by L. A. Zadeh*, John Wiley & Sons, pp.467-498, 1987.

[13] H.-J. Zimmermann, *Fuzzy set theory and Its applications*, Kluwer-Nijhoff Publishing, 1986.

[14] J. Zhang and A. Knoll, "Designing fuzzy controllers by rapid learning," *Fuzzy sets and systems*, Vol.101, pp.287-301, 1999.

[15] R. Zwick et al, "Measures of Similarity among Fuzzy Concepts : A Comparative Analysis," International Journal of Approximate Reasoning, Vol.1, pp.221-242, 1987.

[16] R. Degani and G. Bortolan, "The Problem of Linguistic Approximation in Clinical Decision Making," Vol.2, pp.143-162, 1988.

최 대 영

e-mail : dychoi@green.yuhan.ac.kr

1985년 서강대학교 컴퓨터학과 졸업(학사)

1985년~1990년 한국국방연구원(KIDA) 전산센타 연구원

1992년 서강대학교 이공 대학원 컴퓨터학과 졸업(석사)

1994년 정보처리기술사(전자계산 조직응용 분야)

1996년 서강대학교 이공 대학원 컴퓨터학과 졸업(박사)

1997년~현재 유한대학 경영정보과 조교수

관심분야 : Business Intelligence(BI), 지식 표현, 데이터 마이닝